

PENGUNAAN SUPPORT VECTOR REGRESSION (SVR) PADA PREDIKSI RETURN SAHAM SYARIAH BEI

Retno Maharesi

*Fakultas Teknologi Industri Jurusan Teknik Informatika Universitas Gunadarma
Jl. Margonda Raya 100. Pondokcina Depok
rmaharesi@staff.gunadarma.ac.id*

Abstrak

Pada artikel ini algoritma support vector regression (SVR) digunakan untuk mendapatkan model prediksi return saham syariah di bursa efek Indonesia. Sampel adalah emiten saham dengan likuiditas tinggi selama periode 2012. Pada penelitian ini Pembentukan model didasarkan pada sebuah persamaan yang menghubungkan nilai PBV dan ROE. Variabel terikat pada model adalah nilai proporsi rerata tahunan harga saham pada dua tahun berurutan. Data harga saham merupakan hasil perkalian Price to book value (PBV) dan Book value (BV). Sedangkan variabel bebasnya terdiri atas Book value (BV), tingkat pengembalian ekuitas (ROE) dan proporsi deviden yang dibayarkan ke investor public (POR). Performansi model prediksi berbasis Support Vector Machine (SVR) selanjutnya dibandingkan dengan model Regresi linear berganda berbasis Ordinary Least Squares (RLB-OLS) menggunakan pengukuran nilai Mean square error dan korelasi kuadratik untuk kesesuaian model. Hasil perbandingan kedua model memperlihatkan bahwa model prediksi yang didapat menggunakan model SVR lebih baik.

Kata Kunci: *regresi linear, model prediksi, mean square error, algoritma support vector machine.*

PENDAHULUAN

Apabila prinsip berinvestasi dijalankan secara benar peluang mendapat hasil investasi di pasar modal menjadi lebih besar, hal ini dapat diamati dengan munculnya banyak perusahaan manajemen investasi yang beroperasi di beberapa bank Nasional untuk memasarkan produk reksadannya. Sebuah aktifitas penting terkait dengan investasi di pasar modal adalah kegiatan analisis harga saham, yang bertujuan mengestimasi tingkat harga wajar suatu saham sehingga dapat diperkirakan besar peluang suatu saham mengalami peningkatan harga, yang pada gilirannya memberikan keuntungan bagi pembeli efek tersebut. Pada instrument investasi saham salah satu komponen penentu

tingkat pengembalian adalah *gain* yaitu selisih positif antara harga jual dengan harga beli.

Terdapat dua faktor yang mempengaruhi harga suatu saham yaitu faktor internal dan faktor eksternal (Zainuddin dan J.Hartono, 1999; Triyono dan J. Hartono, 2000; N. F Asyik dan Soelistyo, 2000). Faktor internal adalah faktor dari kinerja perusahaan yang dapat diketahui melalui laporan keuangan perusahaan. Rasio yang digunakan biasanya diukur dari lima rasio keuangannya yaitu rasio likuiditas, solvabilitas, aktivitas, profitabilitas dan saham. Faktor eksternal adalah faktor lingkungan di luar perusahaan yang secara langsung maupun tidak langsung mempengaruhi harga saham. Faktor lingkungan eksternal meliputi kondisi

ekonomi, sosial, politik dan keamanan serta teknologi. Banyak literatur menyatakan bahwa pergerakan harga saham secara umum tergantung pada kondisi ekonomi seperti: Kebijakan moneter, dalam hal ini akan dilihat dari jumlah uang beredar, tingkat suku bunga, kebijakan fiskal, yaitu kebijakan yang berkaitan dengan pajak, jumlah pengeluaran pemerintah (subsidi). Faktor ekonomi lain yang berpengaruh adalah inflasi, Gross National Product (GNP), jumlah pengangguran, produksi nasional, pengeluaran rata-rata masyarakat, investasi dalam bisnis. Selain tiga faktor di atas, yang memengaruhi fluktuasi harga saham adalah kinerja saham, yang merupakan salah satu ukuran dari kemampuan kinerja pasar. Kinerja saham merupakan salah satu faktor pertimbangan untuk menentukan saham pilihan bagi investor. Kinerja saham dapat diukur dari rasio: *Earning per share*, *dividen yield*, *price to earning ratio*, *asset value per share* dan *book value per share*

Model SVR mempunyai keunggulan dibanding model Regresi Linear Berganda (RLB) berbasis estimasi Least Squares, dalam hal pemanfaatan model *nonlinear* secara implisit melalui penerapan fungsi kernel yang memetakan vektor titik data fitur x ke ruang berdimensi lebih tinggi sehingga memungkinkan penggunaan model seperti pada kasus data *linearly separable*. Pada model RLB, hal ini dicapai menggunakan transformasi variabel terikat dan bebas, disesuaikan dengan asumsi linearitas model regresi. Namun seringkali tidak dijumpai fungsi transformasi yang sesuai untuk data. Selain itu, proses training parameter model SVM menggunakan formulasi permasalahan optimasi kuadratik di mana terdapat solusi optimum tunggal. Sehingga model SVM diharapkan terhindar dari permasalahan *overfitting model* dibandingkan dengan model proses training parameter model

Jaringan Syaraf Tiruan (JST) yang kriteria terminasi proses training tidak terdefinisi eksplisit.

Tujuan Penelitian

Pada penelitian ini dikembangkan sebuah model untuk memprediksi tingkat *gain* saham yang terdapat dalam indeks Syariah BEI menggunakan algoritma Support Vector Machine (SVM) khususnya model Support Vector Regression (SVR) untuk meramalkan tingkat pengembalian saham syariah. Permasalahan penelitian adalah bagaimana pemodelan data berbasis algoritma SVR dapat digunakan pada permasalahan prediksi return saham syariah kemudian membandingkan hasil prediksi model SVR dengan hasil prediksi dengan regresi linier berganda.

METODOLOGI PENELITIAN

Metode penelitian adalah sebagai berikut: (1) Sampel emiten syariah diambil secara *purposif sampling* dengan kriteria: Termasuk dalam kategori saham syariah pada tahun yang diteliti, laporan keuangan lengkap sampai tutup buku bulan Desember. Semua data tersedia dan termasuk saham yang aktif dalam perdagangan di pasar modal. Berdasarkan ketentuan ini dari 302 perusahaan yang termasuk dalam kategori syariah terpilih data dari 88 dan 46 perusahaan untuk digunakan pada analisis korelasi dan model prediksi return saham. (2) Melakukan observasi terhadap nilai korelasi *bivariate* dari pasangan variabel yang berasal dari data rasio pengukur kinerja emiten saham syariah untuk mendapatkan gambaran hubungan antara variabel secara linear. (3) Menentukan persamaan matematis yang terdapat dalam literatur sehubungan dengan nilai korelasi maksimal antar variabel rasio pengukur kinerja saham dengan langkah 2. (4) Menentukan model prediksi nilai

variabel terikat berdasarkan nilai variabel-variabel bebas menggunakan model RLB dan SVR. (5) Membandingkan akurasi model yang didapat menggunakan data *training* dan atau data uji. Kelima langkah tersebut dijelaskan secara sebagai berikut.

Data pada penelitian ini diperoleh melalui situs *web* Bursa Efek Indonesia yang berupa *Balance sheet*, *Income Statement*, *Trading activities* dan *Financial Ratios* dari para emiten saham Syariah dalam kurun waktu 2010-2012. Variabel terikat dinyatakan dalam bentuk proporsi dari harga saham pada tahun t terhadap tahun $t-1$ dan variabel bebas terdiri atas 3 variabel prediktor yang meliputi: *ROE (Return on Equity)*, *BV(Book value per share)*, *POR (Price to earning ratio)*. Estimasi harga digunakan dalam variabel terikat didapat melalui hasil perkalian BV terhadap data PBV (*Price to book value*). Penentuan awal model prediksi tingkat pengembalian saham syariah dijalankan menggunakan analisis korelasi data kinerja 88 saham syariah (berupa tujuh macam rasio keuangan). Kemudian dicari set variabel dengan nilai korelasi *bivariate* yang memberikan nilai korelasi relatif tinggi antar variabel dalam kelompok sebagai pertimbangan awal dalam menentukan model prediksi harga saham.

Regresi Linear Berganda

Analisis regresi adalah teknik statistik untuk pemodelan atau mengetahui adanya hubungan fungsional antara satu variabel terikat dengan satu atau lebih variabel bebas. Persamaan regresi linier berganda dengan n data sampel berupa bidang *hyperplane* jika jumlah variabel bebas $m > 2$ dinyatakan dengan Persamaan (1):

$$y_i = b + b_1x_1 + b_2x_2 + \dots + b_mx_m + e_i \\ i = 1, 2, \dots, n. \quad (1)$$

Persamaan (1) dapat dinyatakan dalam notasi vektor dengan vektor galat e berukuran $n \times 1$, vektor baris koefisien variabel bebas w^T berukuran $1 \times m$, vektor variabel bebas x berukuran $m \times 1$ di mana $e = y - w^T X$. Nilai estimasi untuk w didapat dengan meminimalkan Jumlah Kuadrat Galat (JKG). Solusi tunggal dicapai asalkan nilai determinan dari sistem tidak nol. Persamaan regresi penduga yang didapat dengan menghilangkan komponen galat e_i pada Persamaan (1) jika ditulis dalam notasi matriks X ($n \times m$), y ($n \times 1$), w ($m \times 1$) memberikan Persamaan (2)

$$y = w^T X \quad (2)$$

Persamaan regresi penduga (2) diperoleh dengan menyelesaikan sebuah sistem persamaan linear dalam bentuk matriks untuk vektor koefisien w asalkan matriks $(X^T X)$ singular atau determinannya tidak sama dengan nol, yaitu:

$$\hat{w} = (X^T X)^{-1} X^T y.$$

Serangkaian pengujian statistika diperlukan untuk dapat meloloskan model seperti pada Persamaan (1) dapat digunakan dalam proses prediksi suatu nilai variabel terikat.

Seringkali untuk keperluan pembentukan model berbasis algoritma jaringan seperti Jaringan syaraf tiruan (JST) dan *Support Vector Machine* (SVM), nilai *intercept* b pada model regresi ditulis secara terpisah dari suku lain. Hal ini karena nilai tersebut merepresentasikan bias pada model. Sehingga menggunakan alasan tersebut jumlah kuadrat terkecil yang ditulis dalam bentuk perkalian titik vektor,

$w = (b, b_1, b_2, \dots, b_m)^T$, dan $x_i = (x_{i1}, x_{i2}, \dots, x_{im})^T$ adalah:

$$\epsilon^T \epsilon = \min \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - (w \cdot x_i + b))^2 \quad (3)$$

SUPPORT VECTOR REGRESSION (SVR)

Model SVR didasarkan pada pengembangan algoritma *support vector machine* (SVM) yang memuat komponen *noise* dan nonlinearitas yang semula digunakan untuk menyelesaikan permasalahan klasifikasi dua kelas berbeda. Pada algoritma SVR, variabel terikat mempunyai jangkauan nilai berupa bilangan real atau $y_i \in \mathbb{R}^1$. Misalnya diperoleh data training (x_i, y_i) di mana $i = 1, 2, \dots, n$ dengan input $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ dengan $x_i \in \mathbb{R}^m$ dan $y_i \in \mathbb{R}^1$. Sebagai pengingat, pada model pengklasifikasi SVM, data fitur $x_i^T = (x_{i1}, x_{i2}, \dots, x_{im})$ diklasifikasi atas dua kelas (positif atau negatif) menggunakan variabel terikat $y_i \in \{-1, 1\}$ berdasarkan fungsi keputusan berupa persamaan bidang datar atau *hyperplane* H (jika jumlah atribut $m > 3$). *Hyperplane* H tersebut secara maksimal memisahkan dua kelas berbeda dalam himpunan n data fitur x_i sedemikian hingga jarak titik data dalam himpunan support vektor S maksimal terhadap H . Untuk model regresi, permasalahannya adalah bagaimana menentukan *hyperplane* $f(x_i, \theta) = w \cdot x_i + b$ yang

memenuhi batasan nilai ϵ yang ditentukan sehingga $(y_i - \hat{y}_i) \leq \epsilon$.

Penerapan algoritma SVM untuk sebuah model regresi atau klasifikasi memerlukan *parameter set up* atau dikenal juga dengan istilah *hyperparameter* (parameter yang tidak muncul pada persamaan penduga). Sama seperti pada model pengklasifikasi, persamaan *hyperplane* regresi diperoleh dengan mendapatkan vektor parameter w sedemikian hingga diperoleh sebanyak mungkin titik data vektor dengan jarak titik data terhadap *hyperplane* (tegak lurus pada fungsi penduga $f(x_i, \theta)$ dengan vektor normal w) maksimal sama dengan ϵ . Berikut ini adalah formulasi dual dari model regresi SVM nonlinear mengikuti (Hsu dkk, 2010); Burges C., (1998))

Pada formulasi di atas *hyperparameter* C bertindak sebagai nilai *penalty* untuk setiap nilai galat ξ . *Hyperparameter* yang terdiri atas C dan parameter dari fungsi kernel $K\langle x_i, x_j \rangle$ dengan $K\langle x_i, x_j \rangle = \phi^T(x_i)\phi(x_j)$ ditentukan sebelumnya menggunakan prosedur validasi silang (Andrew Ng, 2009; Cawley dan Talbot, 2010). Pada formulasi di atas *hyperparameter* C bertindak sebagai nilai *penalty* untuk setiap nilai galat ξ .

maksimumkan :

$$-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_{1i} - \alpha_{2i})(\alpha_{1j} - \alpha_{2j}) K\langle x_i, x_j \rangle + \sum_{i=1}^n (\alpha_{1i} - \alpha_{2i}) y_i - \epsilon \sum_{i=1}^n (\alpha_{1i} + \alpha_{2i})$$

terhadap:

$$\sum_{i=1}^n (\alpha_{1i} - \alpha_{2i}) = 0$$

$$0 \leq \alpha_{1i} \leq C, i = 1, 2, \dots, n$$

$$0 \leq \alpha_{2i} \leq C, i = 1, 2, \dots, n$$

(4)

Hyperparameter yang terdiri atas C dan parameter dari fungsi kernel $K(x_i, x_j)$ dengan $K(x_i, x_j) = \phi^T(x_i)\phi(x_j)$ ditentukan sebelumnya menggunakan prosedur validasi silang. Pada ruas kanan persamaan, $\phi(x)$ menunjukkan suatu titik di dalam ruang fitur F hasil pemetaan x di dalam fitur ruang input. Faktor pengali *Langrange* beserta kondisi optimalitas KKT (Karush Kuhn Tucker) menghasilkan fungsi regresi secara eksplisit yaitu:

$$f(x) = \sum_{i=1}^{\lambda} (\alpha_{1i} - \alpha_{2i}) K(x_i, x) + b. \quad (5)$$

Himpunan data *support vector* S diperoleh dengan cara mendapatkan data dengan nilai indeks berkorespondensi dengan $0 < \alpha_{1i}, \alpha_{2i} < C$ dan ξ , sehingga didapatkan:

$$b = y_s - \epsilon - \sum_{i \in S} (\alpha_{1i} - \alpha_{2i}) x_i \cdot x_s,$$

dengan alasan yang sama pada kasus model klasifikasi, estimasi b diperoleh dengan nilai rerata berdasarkan data *support vector*:

$$b = \frac{1}{n_s} \sum_{s \in S} \left(y_s - \epsilon - \sum_{i \in S} (\alpha_{1i} - \alpha_{2i}) x_i \cdot x_s \right)$$

Konstanta $C > 0$ pada formulasi di atas merupakan *penalty* akibat pelanggaran toleransi yang berupa batas atas deviasi ϵ terhadap fungsi $f(x)$. Dalam SVR, *support vector* adalah data training yang terletak pada dan di luar batas fungsi keputusan $|y - f(x)|$ sehingga ukuran jumlah *support vector* yang ditentukan oleh besarnya nilai ϵ menentukan performansi dari model SVR. Sehingga dalam SVR akurasi dari aproksimasi terhadap data *training* ditentukan oleh ϵ . Jika nilai ϵ rendah dapat diharapkan semakin banyak data yang terletak pada batas tabung berdiameter $|y - f(x)| = \epsilon$. Sebaliknya, nilai yang tinggi untuk ϵ berkaitan dengan

semakin sedikitnya data yang terletak pada batas tabung.

Ukuran kebaikan model prediksi

Jika y_i menyatakan nilai pengamatan data ke- i dan \hat{y}_i adalah nilai prediksi untuk data ke- i berdasarkan n data sampel maka ukuran galat untuk menilai ketepatan model pada prosedur validasi estimasi model menggunakan data sampel adalah *Mean squared error* (MSE) dan R^2 (Walpole, 1995) sebagai berikut:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

(6)

dan

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n-1) \sum_{i=1}^n (y_i - \bar{y})^2}$$

(7)

Berdasarkan ukuran galat di atas, dapat diketahui model yang sesuai, yaitu jika kecil nilai MSE kecil dan proporsi variabilitas yang dijelaskan oleh model (R^2) besar. Sehingga dapat ditentukan model yang dapat digunakan untuk memprediksi nilai variabel terikat.

HASIL DAN PEMBAHASAN

Dari analisis korelasi terdapat korelasi tinggi antara PBV dan ROE . ROE adalah pengukur kemampuan saham dalam memberikan *return* kepada pemegang saham publik. ROE lebih berperan dalam keputusan pembelian dibandingkan dengan ROA yang mengukur kemampuan asset perusahaan dalam memberikan *return* kepada kreditor dan pemegang saham preferen (Hanafi dan Halim, 2012). Untuk kasus saham syariah Indonesia hal ini dapat dibuktikan oleh hasil regresi sebagai berikut: $PBV = 1.093 - 0.361 ROA +$

0.342 ROE - 1.771 DER, pada persamaan regresi penduga terlihat terdapat korelasi negatif antara ROA dan PBV.

Model deviden terdiskon pada buku (Hanafi dan Halim, 2012) digunakan untuk memperkirakan harga wajar per lembar saham saat melakukan pembelian. Deviden diestimasi menggunakan rumus ROE dikalikan rasio deviden yang dibayarkan ke pemegang saham publik (POR). Sehingga dengan demikian harga *present value* P_0 untuk ekspektasi harga pada periode berikutnya untuk tingkat pertumbuhan g dan laju diskon untuk saham tertentu r , dapat dituliskan dengan Persamaan 8.

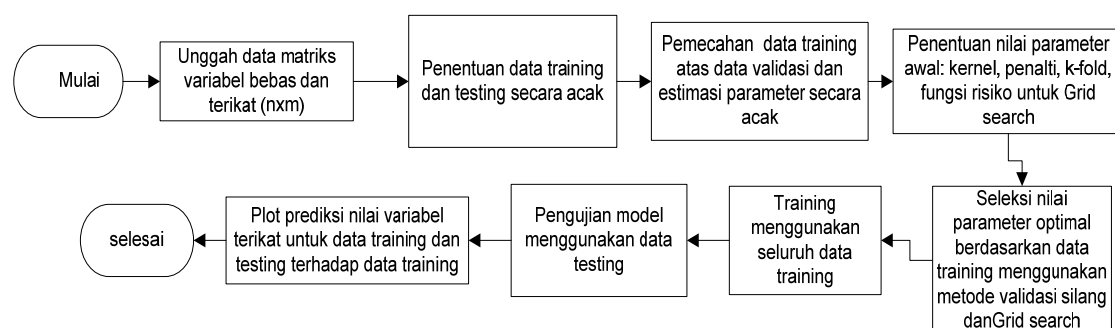
$$P_0 = \frac{ROE \cdot POR}{r - g} \quad (8)$$

Berdasarkan Persamaan (8) dan penggunaan rasio pasar PBV, yaitu $PBV = P_0/BV$ dapat dibentuk fungsional dari model prediksi untuk prakiraan harga per unit saham yaitu:

$$P_t = f(BV, ROE, POR) = a(BV_t \cdot ROE_t \cdot POR_t)$$

Data training dan uji secara keseluruhan terdiri atas 46 observasi dengan variabel terikat adalah rasio estimasi harga per unit saham pada tahun 2011 dan 2010 sedangkan variabel bebas adalah proporsi nilai BV, ROE dan POR untuk dua periode berurutan.

Algoritma pembelajaran supervisi SVR membutuhkan data *training* sampel. Oleh karena pada model SVR juga terdapat Parameter *set up*, maka untuk proses seleksi model digunakan metode validasi silang *k-lipatan* untuk menghasilkan model terbaik dengan nilai parameter optimal. Proses pencarian parameter model dan *set up parameter* (*hyperparameter*) menggunakan metode *grid search*. Proses seleksi model SVR dilakukan melalui beberapa eksperimen menggunakan fungsi kernel Radial Basis, linear dan polinom. Parameter fungsi kernel terbaik diperoleh dengan metode *grid search* kasar yang dijalankan pada interval pencarian 2^{-7} sampai dengan 2^7 menggunakan interval perubahan 0.2. Batas atas nilai *penalty* menggunakan selisih nilai maksimal dan minimal data variabel terikat. Batas toleransi ϵ diambil nilai antara 0 sampai dengan 5 menggunakan interval perubahan 0.1. Untuk ukuran performansi model prediksi digunakan ukuran kesalahan MSE dan ukuran kesesuaian model dengan data yaitu R^2 . Gambar 1 adalah diagram alur proses seleksi model SVR. Model SVR yang dihasilkan diperoleh setelah menjalankan skript berulang kali, dan didapatkan hasil yang stabil untuk berbagai percobaan menggunakan nilai interval parameter yang berbeda.



Gambar 1. Diagram alur proses seleksi model SVR

Tabel 1.
Ringkasan hasil Percobaan model SVR dan RLB

Model	Nilai set up hyperparameter				Ringkasan karakteristik Model		
	Jenis Kernel	Interval Epsilon	Interval Gamma	penalty	Jumlah Support Vector	MSE	R ² Training/Uji
SVR	Radial	[0:5]	2[-7 : 7]	Selisih	39	0.0916	0.6934/ 6182
SVR	Polinom-3	dengan interval	dengan interval	nilai max dan	33	0.2079	0.2894/ 0.2844
SVR	Polinom-2	perubahan	perubahan	min	33	0.2894	0.2079
SVR	Linear	0.1	0.2	variabel terikat	22	0.191	0.3460/ 0.3514
OLS ln	x	x	x	x	x	0.141	0.6350/-
OLS	x	x	x	x	x	0.278	0.0210/-

Untuk memperbandingkan kinerja algoritma SVR dalam menghasilkan model prediksi, digunakan regresi linear berganda. Model regresi linear berganda didapat menggunakan data yang sama dengan model SVR dan data yang sudah ditransformasi dengan fungsi logaritma natural. Model RLB yang ditampilkan lolos uji koefisien regresi maupun Anova. Tabel 1 berisi ringkasan hasil estimasi model Regresi Linear Berganda dengan SPSS 17 dan SVR. Model SVR didapat dengan menjalankan skript Matlab sesuai dengan diagram alir proses pada Gambar 1 menurut jenis fungsi kernelnya menggunakan *tool box* LIBSVM 1.73.

Analysis Model Prediksi Return Saham

Dari Tabel 1, terlihat pada kelompok model SVR, nilai MSE terkecil dan R² terbesar dihasilkan oleh eksperimen dengan fungsi kernel Radial dengan besar secara berurutan 0.0916 dan 0.6934. Proses seleksi parameter optimal dilakukan menggunakan metode validasi silang 5 lipatan dengan set up parameter awal seperti terdapat pada Tabel 1. Kami juga mencoba untuk menjalankan skript menggunakan jumlah lipatan berbeda, yaitu 10, namun hasil di atas relatif tidak berubah dari sisi performansi.

Pada tabel 1 kolom terakhir, terlihat performansi model SVR berdasarkan nilai

R² pada pada data uji mengalami penurunan untuk fungsi kernel Radial, namun untuk fungsi kernel polinom maupun linear relatif tetap. Pada Tabel 1 juga terlihat untuk model regresi linier berganda dengan OLS, nilai R² terbesar dicapai setelah melakukan transformasi logaritma natural pada data yang sama seperti pada model SVR dengan persamaan regresi penduga: $\ln(P11/10) = -0.193 + 1.029\ln(BV11/10) + 0.344\ln(ROE11/10) + 0.179\ln(POR11/10)$. Sedangkan pada data yang tidak melalui transformasi logaritma natural diperoleh hasil yang tidak signifikan untuk meloloskan model prediksi. Berdasarkan Tabel 1, perbandingan antara model prediksi OLS dan model SVR memberikan nilai R² /MSE yang lebih baik untuk model SVR menggunakan kernel Radial. Di mana untuk mencapai hasil yang kurang lebih sama, model RLB – OLS memerlukan penerapan transformasi fungsi logaritma sesuai dengan persamaan: $P_t = \alpha (BV_t * ROE_t * POR_t)$. Salah satu metode untuk memperkecil fluktuasi nilai data adalah dengan menggunakan variabel berupa nilai rasio, dalam hal ini digunakan variabel rasio dua nilai pada periode yang berurutan $t= 2011$ dan $t-1= 2010$. Sehingga diperoleh:

$$\frac{P_t}{P_{t-1}} = \alpha_0 \left(\frac{BV_t}{BV_{t-1}} * \frac{ROE_t}{ROE_{t-1}} * \frac{POR_t}{POR_{t-1}} \right) \quad (9)$$

Pada aplikasi model SVR untuk peramalan harga, nilai variabel terikat yang lebih besar dari pada 1 menunjukkan nilai harapan untuk saham dengan nilai variabel bebas diketahui mengalami kenaikan harga dan jika nilainya diantara 0 dan 1 maka berpeluang untuk mengalami penurunan harga. Penerapan transformasi logaritma natural pada data yang diperoleh menggunakan Persamaan (9) memungkinkan untuk menggunakan prosedur estimasi model Regresi Linear Berganda sebagai berikut:

$$\ln \left(\frac{P_t}{P_{t-1}} \right) = \ln \left(\alpha_0 \left(\frac{BV_t}{BV_{t-1}} * \frac{ROE_t}{ROE_{t-1}} * \frac{POR_t}{POR_{t-1}} \right) \right) = \\ a_0 + a_1 \ln \left(\frac{BV_t}{BV_{t-1}} \right) + a_2 \ln \left(\frac{ROE_t}{ROE_{t-1}} \right) + a_3 \ln \left(\frac{POR_t}{POR_{t-1}} \right)$$

Model prediksi *return* saham menggunakan algoritma *Support Vector Regression* (SVR) lebih baik dibandingkan dengan model regresi linear OLS. Model OLS menjadi lebih baik setelah menerapkan transformasi logaritma natural pada data. Hal ini membuktikan Bahwa model prediksi *return* saham adalah *nonlinear*. *Non-linearitas* data dapat berasal dari faktor kinerja internal yang juga dipengaruhi oleh tingkat penjualan di mana kondisi eksternal seperti nilai tukar valuta asing, tingkat suku bunga bank sentral, kebijakan fiskal pemerintah turut serta menentukan tingkat permintaan dan penawaran; sehingga berakibat pada fluktuasi harga saham.

SIMPULAN dan SARAN

Model prediksi SVR terbaik berdasarkan data training menggunakan fungsi kernel Radial Basis karena memiliki nilai *MSE* terendah dan *R*² terbesar dibandingkan dengan peng-

gunaan kernel fungsi polinomial berderajat satu sampai dengan lima.

Pada penelitian ini data yang digunakan adalah data *cross sectional* dalam satu tahun. Oleh karena sifat nonlineartitas data perubahan harga saham sangat tinggi, maka disarankan digunakan data yang terbaru dalam interval waktu yang lebih pendek. Data kuartalan atau semesteran sebaiknya digunakan untuk lebih dapat mengantisipasi struktur nonlineartitas pada perubahan *return* saham menggunakan formulasi model nonlinear SVR pada proses peramalan sehingga hasil prediksi memiliki akurasi yang baik yaitu nilai *R*² tinggi untuk data uji.

DAFTAR PUSTAKA

- Andrew, N. 2001. *Lecture notes: Support Vector Machines*, USA.
- Asyik, N.F., & Soelistyo. 2000. Kemampuan rasio keuangan dalam memprediksi laba, penetapan rasio keuangan sebagai discriminator. *Jurnal Ekonomi dan Bisnis Indonesia*, 15, 313-33.
- Burges C. 1998. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121-167,
- Cawley, G.C., & Talbot N.L.C. 2010. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11. 2079-2107.
- Hanafi, M.H., & Halim, A. 2012. Analisis Laporan Keuangan, Edisi 4 ISBN:979-8170-43-1, STIM YKPN, Yogyakarta.
- Hsu, C.W., Chang, C.C., & Lin, C.J. 2010. *Lectures notes: A practical guide to support vector classification*, Department of Computer Science National Taiwan University, Taipei 106, Taiwan.
- Triyono, & Hartono J. 2000. Hubungan kandungan informasi arus kas, kom-

- ponen arus kas dan laba akuntansi dengan harga atau return saham, *Jurnal Riset Akuntansi Indonesia*, 3, 54-68.
- Walpole, R.E. 1995. *Pengantar statistika, Edisi 3*. PT Gramedia Pustaka Utama, Jakarta.
- Zainuddin, & Hartono J. 1999. Manfaat rasio keuangan dalam memprediksi pertumbuhan laba: Suatu studi empiris pada perusahaan perbankan yang terdaftar di Bursa Efek Jakarta. *Jurnal Riset Akuntansi Indonesia*, 2, 66-90.